



Australian Bureau of Statistics

2085.0 - Microdata: Australian Census Longitudinal Dataset with Social Security and Related Information, experimental statistics, 2006-2011

Latest ISSUE Released at 11:30 AM (CANBERRA TIME) 23/08/2017 First Issue

Summary

Introduction

INTRODUCTION
AVAILABLE PRODUCTS
APPLYING FOR ACCESS
SUPPORT

INTRODUCTION

In August 2017, information from the Department of Social Services' Social Security and Related Information (SSRI) dataset was added to the Australian Census Longitudinal Dataset (ACLD) to create a new dataset (ACLD-SSRI). The linkage of the ACLD with the SSRI data brings together information about the characteristics and circumstances of people who have interacted with the social security system, and has the potential to increase knowledge about a wide range of socio-economic issues facing Australians and their families.

The ACLD-SSRI file includes a cross sectional sample of the SSRI data collected including all individual recipients and partner information where available as of the end of September 2011. Data from this period was selected to best align with the 2011 Census.

While every effort has been made to assure the quality of the statistics presented in this publication, they should be considered experimental and treated with caution. Refer to Limitations for further information.

SSRI data is collected by the Australian Government Department of Human Services on behalf of the Australian Government Department of Social Services for the delivery of Pensions and Payments under the *Social Security Act 1991*, *A New Tax System (Family Assistance) Act 1999* and the *Student Assistance Act 1973*.

The Census of Population and Housing is conducted every five years and aims to measure accurately the number of people and dwellings in Australia on Census Night. The Census also provides information on the key characteristics of people and dwellings for small geographic areas and for small population groups.

The Australian Census Longitudinal Dataset (ACLD) brings together a 5% sample from the 2006 Census with corresponding records from the 2011 Census to create a research tool for exploring how Australian society is changing over time. In taking a longitudinal view of Australian society, the ACLD can uncover new insights into the dynamics and transitions

that drive social and economic change over time, and how these vary for diverse population groups and geographies. Information relating to migrants from the Department of Social Services' (DSS) Settlement Database is included on the ACLD, drawing from the existing data linkages in the Microdata: Australian Census and Migrants Integrated Dataset, 2011 (cat. no. 3417.0.55.001).

The ACLD and the ACLD-SSRI are released available for use in TableBuilder and as microdata products in DataLab.

Microdata products are the most detailed information available from a Census or survey and usually consist of data items derived from individual responses to a questionnaire. ABS microdata is released with the approval of the Australian Statistician under the *Census and Statistics Act 1905*. No microdata can be released in a manner likely to allow an individual to be identified.

Our internal processes ensure that privacy is a paramount consideration when data is being combined.

No person will ever be able to see all of your information together at any point of the data combining process. This is known as the separation principle. Under the Census and Statistics Act 1905, the personal information you provide the ABS remains strictly confidential. The ABS never has and never will release identifiable data. As an Australian Government agency, we also comply with the Privacy Act 1988 and handle your personal information in accordance with the Australian Privacy Principles.

AVAILABLE PRODUCTS

The following microdata products are available for this longitudinal dataset:

- ACLD-SSRI in TableBuilder - an online tool for creating tables and graphs.
- ACLD-SSRI in DataLab - for in-depth analysis using a range of statistical software packages.

Further information about TableBuilder and DataLab, and other information to assist users in understanding and accessing microdata in general, is available from the Microdata Entry Page. Before applying for access to TableBuilder, users should read and familiarise themselves with the information contained in the TableBuilder, User Guide (cat. no. 1406.0.55.005)

APPLYING FOR ACCESS

To apply for access to TableBuilder, please register and apply in the Registration Centre.

Information on TableBuilder access can be found on the How to Apply for Microdata page on the ABS web site.

If you are already a registered TableBuilder user [Login here](#).

To apply for access to the ACLD or ACLD-SSRI in DataLab, please contact Microdata Access Strategies via microdata.access@abs.gov.au.

Information on DataLab can be found on the [About the DataLab](#) page on the ABS web site.

SUPPORT

For support in the use of this product, please contact Microdata Access Strategies on 02 6252 7714 or via microdata.access@abs.gov.au.

DATA AVAILABLE ON REQUEST

Customised tables are available on a fee-for-service basis. A consultancy service is available for complex analysis and modelling. For further information, contact the National Information and Referral Service on 1300 135 070 or email client.services@abs.gov.au.

INQUIRIES

For further information about these and related statistics, contact the National Information and Referral Service on 1300 135 070, or email client.services@abs.gov.au. The ABS Privacy Policy outlines how the ABS will handle any personal information that you provide to us.

Insights from data integration: Australians on the Age Pension

Australians on the **AGE PENSION**

The Age Pension provides income support and access to a range of concessions for eligible women aged 64 years and over, and men aged 65 years and over.



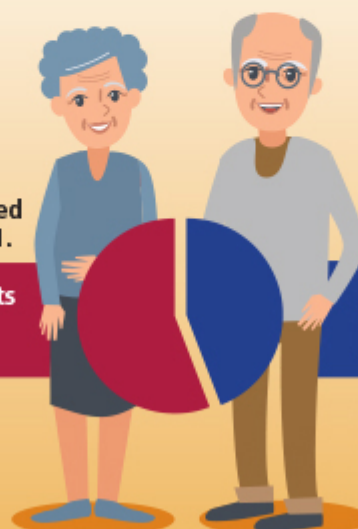
Who received the Age Pension?



69%

of Australians of
pension age received
the pension in 2011.

56% of recipients
were
Women



44% of recipients
were
Men

75
Years

The average
age of a person
on the pension.

How did they spend their time?

Caring



12%

cared for children.

Of those, 61% were women and 39% were men.

Women

Men

61%

39%



12%

cared for someone with a disability.

Volunteering



24%

of 65 year old pensioners did voluntary work.

This reduced to **8%** of those aged 85 and over.



Paid work



6%

worked in some capacity, mostly part-time.

Of those employed the most common paid work for men was Road and Rail Drivers (11%). Women were most commonly Carers and Aides (9%)



11%



9%



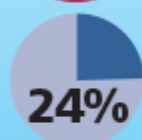
What were their living and housing arrangements?

Living arrangements in 2011



66%

of pensioners lived at the same address as they did 10 years earlier.



24%

of pensioners aged under 80 lived alone.

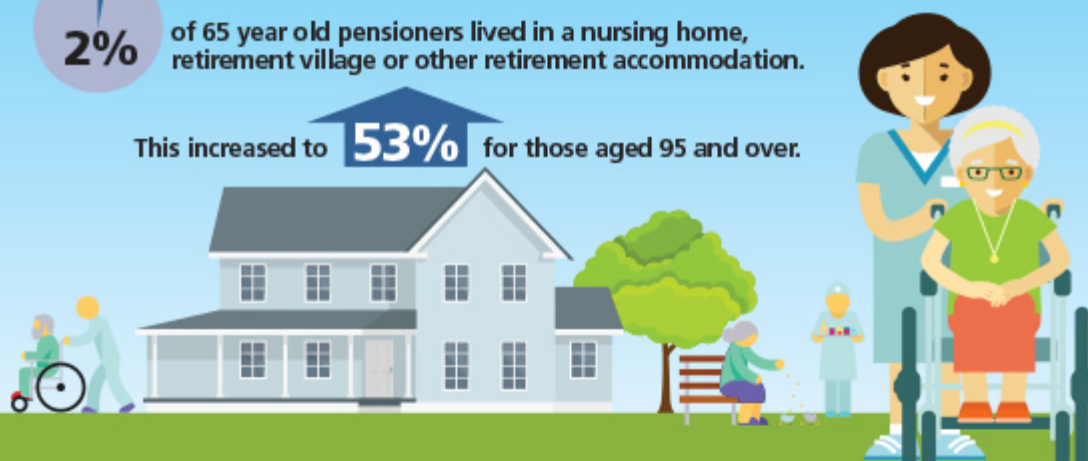
This increased to nearly **40%** for those aged 80 and over.



2%

of 65 year old pensioners lived in a nursing home, retirement village or other retirement accommodation.

This increased to **53%** for those aged 95 and over.



Changes in type of accommodation



8%

changed their type of accommodation from **2006** to **2011**.



Home ownership



75%
owned their
home outright.



8%
still had a
mortgage.



Regional variations



**Pensioners who owned
their home outright**

73%
in Albany

91%
in Keilor

38%
in Sydney
Inner City

Renting



15% were renting, with **35%** of those living in public housing.



Home internet access



52%
had access to the
internet at home.



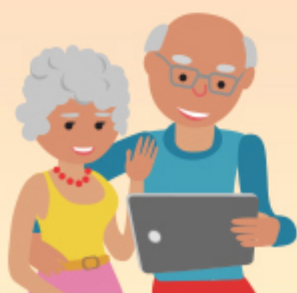
Lowest level
of home internet
access: NT,

47%

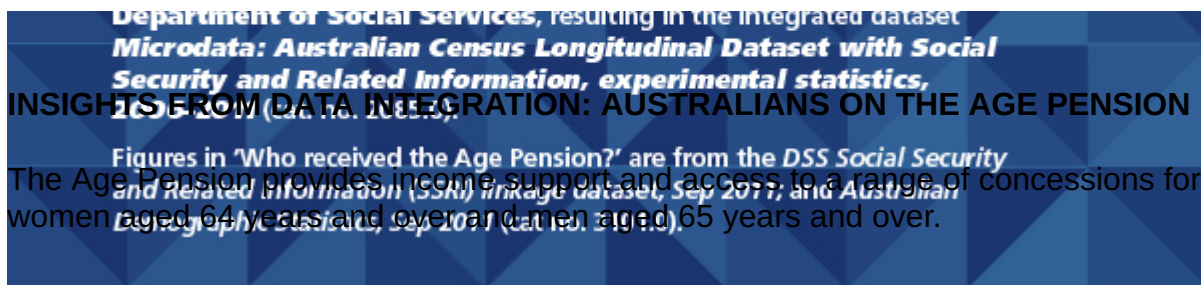


Highest level
of home internet
access: ACT,

65%



Where did these numbers come from?



WHO RECEIVED THE AGE PENSION?

69% of Australians of pension age received the pension in 2011.

56% of recipients were women . 44% of recipients were men.

The average age of a person on the pension is 75 years.

HOW DID THEY SPEND THEIR TIME?

Caring

12% cared for children. Of those, 61% were women and 39% were men.

12% cared for someone with a disability.

Volunteering

24% of 65 year old pensioners did voluntary work. This reduced to 8% of those aged 85 and over.

Paid work

6% worked in some capacity, mostly part-time.

Of those employed, the most common paid work for men was Road and Rail Drivers (11%).

Women were most commonly Carers and Aides (9%).

WHAT WERE THEIR LIVING AND HOUSING ARRANGEMENTS?

Living arrangements in 2011

66% of pensioners lived at the same address as they did 10 years earlier.

24% of pensioners aged under 80 lived alone. This increased to nearly 40% for those aged 80 and over.

2% of 65 year old pensioners lived in a nursing home, retirement village or other retirement accommodation. This increased to 53% for those aged 95 and over.

Changes in type of accommodation

8% changed their type of accommodation from 2006 to 2011. The most common change was to a retirement village from another private dwelling.

Home ownership

75% owned their home outright. 8% still had a mortgage.

Regional variations:

Pensioners who owned their home outright:

91% in Keilor

73% in Albany

38% in Sydney Inner City

Renting

15% were renting, with 35% of those living in public housing.

Home internet access

52% had access to the internet at home.

Highest level of home internet access: ACT, 65%.

Lowest level of home internet access: NT, 47%.

WHERE DID THESE NUMBERS COME FROM?

A joint project between the Australian Bureau of Statistics and the Department of Social Services, resulting in the integrated dataset *Microdata: Australian Census Longitudinal Dataset with Social Security and Related Information, experimental statistics, 2006–2011* (cat. no. 2085.0).

Figures in 'Who received the Age Pension?' are from the *DSS Social Security and Related Information (SSRI) linkage dataset, Sep 2011*; and *Australian Demographic Statistics, Sep 2011* (cat no. 3101.0).

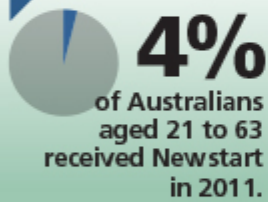
Insights from data integration: Australians on Newstart

Australians on **NEWSTART**

Newstart Allowance provides financial assistance to eligible 21 to 63 year olds who are looking for work or participating in approved activities to increase their chances of finding a job.



Who received Newstart Allowance?



41% of recipients were
Women



59% of recipients were
Men

40
Years

The average
age of people
receiving
Newstart.

Who did they support and care for?

Almost
50%
had dependent
children

18%
were lone parents

14%
cared for a person
with a disability

73%
Mums

27%
Dads

How did housing affect their lives?

Around
50%
rented

Those **renting privately** were more than twice as likely to have changed address in the past year than those in **public housing**.

Of these,
69%
rented privately

and almost
25% were
public housing
tenants

What had they studied?



Almost
40%
had a non-school qualification
compared with 58% of Australians
of the same age.



**Most common
qualifications:**

17%

Certificate Level III & IV
compared with 20% of Australians of the
same age.



7%

had a
Bachelor degree
compared with 18%
of Australians of the
same age.



2%

had a
Postgraduate degree
compared with 7% of
Australians of the
same age.



Study Interests:

People on Newstart were twice as likely
to have studied the **Creative Arts** ...



... but around half as likely to have
studied **Education** compared with
Australians of the same age.



What work were they doing?

Voluntary work in 2011

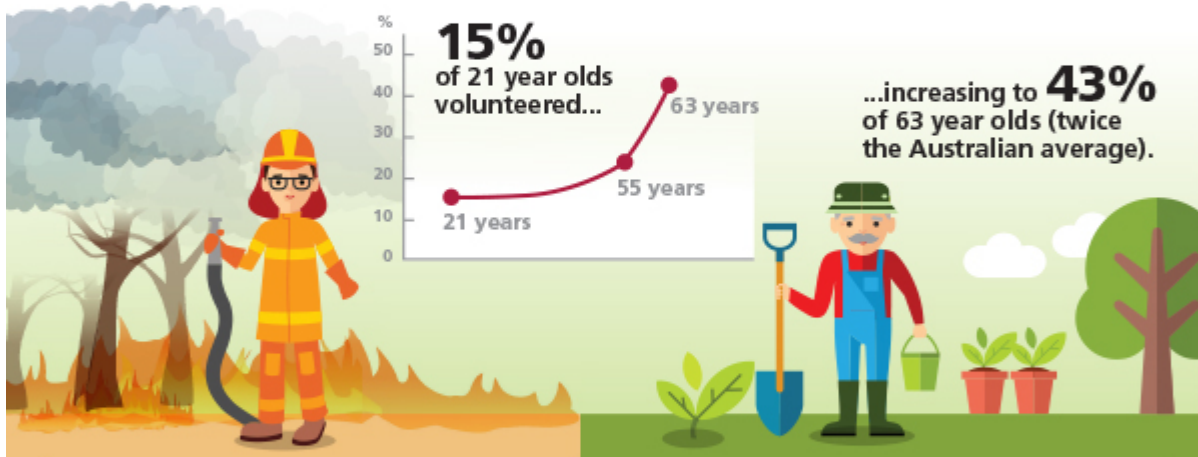
The mutual obligation requirements for people on Newstart encourage those aged 55 or more to do voluntary work and support those under 55 years to volunteer where it will help them to find work.



Overall, **20%** of people on Newstart in 2011 did voluntary work. Of these, 36% had also volunteered in 2006.



Volunteering increased with age:



Paid work in 2006

52% of people on Newstart in 2011 were employed in 2006



The most common occupation for **men** was Factory Process Workers (6%).



LOOKING BACK

Women were most commonly Sales Assistants and Salespersons (14%).



Where did these numbers come from?

A joint project between the **Australian Bureau of Statistics** and the **Department of Social Services**, resulting in the integrated dataset ***Microdata: Australian Census Longitudinal Dataset with Social Security and Related Information, experimental statistics, 2006–2011*** (cat. no. 2085.0).

Figures in 'Who received Newstart Allowance?' are from the *DSS Social Security and Related Information (SSRI) linkage dataset, Sep 2011*; and *Australian Demographic Statistics, Sep 2011* (cat no. 3101.0).

INSIGHTS FROM DATA INTEGRATION: AUSTRALIANS ON NEWSTART

The Newstart Allowance provides financial assistance to eligible 21 to 63 year olds who are looking for work or participating in approved activities to increase their chances of finding a job.

WHO RECEIVED NEWSTART ALLOWANCE?

4% of Australians aged 21 to 63 received Newstart in 2011.
59% of recipients were men. 41% of recipients were women.
The average age of people receiving Newstart: 40 years.

WHO DID THEY SUPPORT AND CARE FOR?

Almost 50% had dependent children.
18% were lone parents (73% mums, 27% dads).
14% cared for a person with a disability.

HOW DID HOUSING AFFECT THEIR LIVES?

Around 50% rented. Of these, 69% rented privately and almost 25% were public housing tenants.
Those renting privately were more than twice as likely to have changed address in the past year than those in public housing.

WHAT HAD THEY STUDIED?

Almost 40% had a non-school qualification (compared with 58% of Australians of the same age).

Most common qualifications:

17%: Certificate Level III & IV (compared with 20% of Australians of the same age).
7% had a Bachelor degree (compared with 18% of Australians of the same age).
2% had a Postgraduate degree (compared with 7% of Australians of the same age).

Study interests:

People on Newstart were twice as likely to have studied the Creative Arts, but around half as likely to have studied Education compared with other Australians of the same age.

WHAT WORK WERE THEY DOING?

Voluntary work in 2011

The mutual obligation requirements for people on Newstart encourage those aged 55 or more to do voluntary work and support those under 55 years to volunteer where it will help them to find work.

Overall, 20% of people on Newstart in 2011 did voluntary work. Of these 36% had also volunteered in 2006.

Volunteering increased with age:

15% of 21 year olds volunteered, increasing to 43% of 63 year olds (twice the Australian average).

Looking back: Paid work in 2006

52% of people on Newstart in 2011 were employed in 2006.

The most common occupation for men was Factory Process Workers (6%).

Women were most commonly Sales Assistants and Salespersons (13%).

WHERE DID THESE NUMBERS COME FROM?

A joint project between the Australian Bureau of Statistics and the Department of Social Services, resulting in the integrated dataset; *'Microdata: Australian Census Longitudinal Dataset with Social Security and Related Information, experimental statistics, 2006–2011* (cat. no. 2085.0).

Figures in "Who received Newstart Allowance?" are from the *DSS Social Security and Related Information (SSRI) linkage dataset, Sep 2011*; and *Australian Demographic Statistics, Sep 2011* (cat no. 3101.0).

Methodology

SCOPE AND COVERAGE

LINKING METHODOLOGY

WEIGHTING, BENCHMARKING AND ESTIMATION

SOURCES OF ERROR

DATA CONSISTENCY

SCOPE AND COVERAGE

The ACLD is a random 5% sample of persons enumerated in Australia on Census Night, 8 August 2006 which has been linked using statistical techniques to records from the 2011 Census, conducted on 9 August 2011. The Census covers all areas in Australia and includes persons living in both private and non-private dwellings, but excludes:

- diplomatic personnel of overseas governments and their families;
- and Australian residents overseas on Census Night.

Overseas visitors to Australia are excluded from the 2006 ACLD sample while persons within Australia who were away from their place of usual residence on Census Night are included.

For more information on the scope and coverage of the Census:

- How Australia Takes a Census, 2006 (cat. no. 2903.0)
- How Australia Takes a Census, 2011 (cat. no. 2903.0)

Additional migrant-related data

There are 18,859 records on the ACLD that have additional data attached resulting from the linkage of 2011 Census data to the Department of Social Services' Settlement Database (SDB). These records correspond to people who had a permanent visa record on the Settlement Database with a date of arrival between January 1, 2000 and August 8, 2006 (that is, Census Night in 2006) and were able to be linked to the ACLD.

The SDB date of arrival on which the scope is based reflects an individual's latest arrival pertaining to their latest permanent visa. For an offshore applicant, the SDB arrival date is when the applicant arrives in Australia on that permanent visa. However, for a person who applies onshore for a permanent visa, the date of arrival listed on the SDB is the date of their last entry into Australia.

For further information about coverage issues, please see Research Paper: Assessing the Quality of Linking Migrant Settlement Records to 2011 Census Data (cat. no. 1351.0.55.043).

Additional social security related data

There are 214,713 records on the ACLD-SSRI that have additional data attached resulting from the linkage of the ACLD to a subset of the Department of Social Services' Social Security and Related Information (SSRI) dataset. This represents those persons who received social security benefits or had suspended benefits as at September 2011 and where their social security record was able to be linked to the ACLD. Data from this period was selected to best align with the 2011 Census. For more detail on the specific benefits available at the time, including benefit eligibility criteria see A Guide to Australian Government Payments produced by the Department of Human Services for September 2011.

Records with ACLD, migrants and social security related data

There are 4,038 records that have both migrant-related data and SSRI data alongside the Census information.

The population these 4,038 records represent corresponds to people who were:

- residents of Australia at the time of and participated in both the 2006 and 2011 Censuses;
- recent migrants of Australia who were granted a permanent visa between January 1, 2000 and August 8, 2006 (that is, Census night in 2006);
- were a social security recipient in September 2011 (that is, they were receiving social security benefits or had suspended benefits as of the end of September); and
- were able to have both their migrant-related record and their social security

record linked to the ACLD.

The records with both migrants and SSRI data are a result of two separate and independent linkage exercises that were performed to create the 2016/2017 ACLD refreshes:

- Linkage of 2011 Census records to DSS Settlement Database records; and
- Linkage of ACLD 2011 Census records to DSS SSRI dataset records.

The ACLD contains migrant information for a subset of the records, that is, those 2011 Census records that were linked to a record in the DSS Settlement Database as a result of linking the 2011 Census to the DSS Settlement Database to produce Australian Census and Migrants Integrated Dataset, 2011 (cat. no. 3417.0.55.001). ACLD 2011 Census records, some of which had previously linked migrant information, were independently linked to records in the SSRI dataset.

While the greatest possible care was taken in both of these linkage exercises, it is not possible to determine the impact of missed links in any analysis of these records with both migration and social security information. Consequently any analysis based on these records should be performed with caution.

LINKING METHODOLOGY

Data from the 2006 ACLD sample and the 2011 Census were brought together using data linkage techniques. The method involved linking without the use of name or address, as this information was destroyed following completion of Census processing from both 2006 and 2011.

Data linkage is typically undertaken using probabilistic and/or deterministic methods, both of which were used in forming the ACLD:

Deterministic linkage involves assigning record pairs across two datasets that match exactly or closely (within specified tolerance levels) on common variables. This type of linkage is most applicable where the records from different sources consistently report sufficient information and can be an efficient process for conducting linkage.

Probabilistic linkage is a method that determines the likelihood that a pair of records are a match based on how well they agree on a set of variables and then uses statistically valid decision rules to designate which record-pairs are matches, possible matches and non-matches. When calculating the likelihood that a pair of records are a match, the discriminatory power of each variable being used for linking is taken into account. This approach also allows links to be assigned in spite of missing or inconsistent information, provided there is enough agreement on other variables to offset any disagreement.

A number of linkage passes were conducted based on different combinations of variables to ensure each record in the random sample taken from the 2006 Census had the highest possible chance of being linked to a record in the 2011 Census. At the end of the linkage process, 800,759 (82%) of the 979,661 records in the sample records from 2006 were linked to a 2011 Census record.

There were two reasons why some records from the 2006 Census were not linked to a 2011 record:

- Records belonging to the same individual were present at both time points but these records failed to be linked because they contained missing or inconsistent information.
- The person had no record in the 2011 Census.

For detailed information on the linking methodology and an assessment of its quality see Australian Census Longitudinal Dataset, Methodology and Quality Assessment (cat. no. 2080.5).

Linkage of migrants related data to the ACLD

The addition of variables relating to migrants, from the Department of Social Services' Settlement Database, to the ACLD is based on the existing linkage in the Australian Census and Migrants Integrated Dataset. For information on the linking methodology of Settlement Database variables see Australian Census and Migrants Integrated Dataset Linking Methodology (cat. no. 3417.0.55.001).

Linkage of social security related data to the ACLD

Variables on the ACLD 2011 Census records and September 2011 SSRI dataset used for linking include:

- Non-identifying grouped numeric code
- Age
- Sex
- Day of birth
- Year of birth
- Country of birth
- Indigenous status
- Marital status
- Meshblock
- Statistical Areas 1 and 2

A 2006 Census Data Enhancement quality study found that in the absence of name and address, inclusion of a non-identifying grouped numeric code based on name improved the accuracy and efficiency of the linking process while still preserving confidentiality. For further information, see Assessing the Likely Quality of the Statistical Longitudinal Census Dataset (cat. no. 1351.0.55.026).

A non-identifying grouped numeric code is an additional linkage item assigned to each record in the dataset. Each numeric code represents approximately 2000 people and therefore is not unique to an individual and cannot be reversed to identify individuals. The grouped numeric codes were created during the Census processing period, and are only accessible to those ABS staff conducting linkages to the ACLD.

In the linkage of the SSRI information to the ACLD the grouped numeric codes were used alongside personal characteristics such as age, sex and meshblock.

A number of linkage passes were conducted based on different combinations of variables to ensure each ACLD 2011 Census record had the highest possible chance of being accurately linked to a record in the DSS SSRI dataset. This resulted in 214,713 links between 2011 Census records on the ACLD and the September 2011 SSRI extract.

As the overlap of these two datasets is unknown, it is difficult to calculate an exact linkage rate for this exercise. However, similar linkage projects conducted by the ABS using many of the same linkage variables have resulted in linkage rates of around 80-85%. The ACLD-SSRI linkage also used a non-identifying grouped numeric code, with which we expect even higher quality results.

There were three reasons why some ACLD 2011 Census records were not linked to a September 2011 SSRI record:

- The person did not either receive social security benefits or have social security benefits suspended as of September 2011.
- The person received social security benefits or had their benefits temporarily suspended as of September 2011, but was not present on the ACLD dataset.
- Records belonging to the same individual were present in both datasets but these records failed to be linked because they contained missing or inconsistent information.

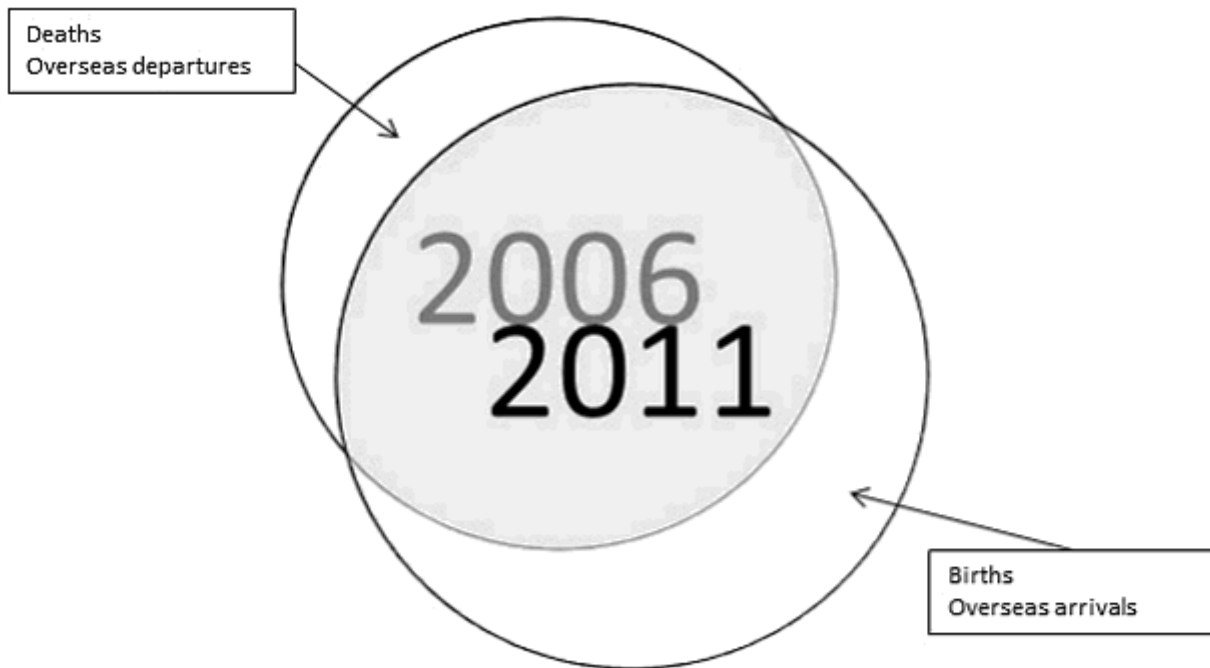
WEIGHTING, BENCHMARKING AND ESTIMATION

Weighting

Weighting is the process of adjusting a sample to infer results for the relevant population. To do this, a 'weight' is allocated to each sample unit - in this case, person. The weight can be considered an indication of how many people in the relevant population are represented by each person in the sample. Weights were created for linked records in the ACLD to enable longitudinal population estimates to be produced. Cross-sectional population estimates for 2006 and 2011 are available from each Census.

The ACLD began as a random sample of 5% of in scope records from the 2006 Census. As such, each person in the sample should represent about 20 people in the population. Between Censuses, however, the in scope population changes as people die or move overseas. In addition, Census net undercount and data quality can affect the capacity to link equivalent records across waves. The ACLD weights benchmarked the linked records to the population that was resident in Australia at the time of both the 2006 and 2011 Censuses. The weights were based on four components: the design weight, undercoverage adjustment, missed link adjustment and population benchmarking.

The original population benchmark was the 2011 Estimated Resident Population (ERP). The 2011 ERP was chosen over the 2006 ERP as the baseline population as it is more recent. The 2011 ERP was then adjusted so as to exclude people who were not in Australia in 2006 as depicted below.



Weights were benchmarked to the following population groups:

- state/territory by age (ten year groups) by sex by mobility (interstate arrivals benchmarked separately).
- Indigenous status by state/territory.

At 12 February 2016 a new weight was applied to the ACLD file to better account for overseas departures and arrivals between 2006 and 2011. Users who have analysed the ACLD prior to 12 February 2016 may notice changes to estimates produced with the revised weight. Estimates of population groups will be different with the total weighted population estimate being 19.5 million compared to 18.6 million on the old weight. Proportions are expected to only show small differences when previous tables are compared.

The weights have a mean value of 24 and range between 17 and 103. Higher weights are associated with people of Aboriginal and Torres Strait Islander origin and people who moved interstate between 2006 and 2011.

While the ACLD weights are believed to be of good quality for use in analysing the longitudinal Australian population and account for missed links between the 2006 and 2011 Censuses, they do not attempt to account for missed links between the ACLD and either the migrant-related data or the social security data.

Estimation

Estimates of population groups are obtained by summing the weights of persons with the characteristic(s) of interest.

SOURCES OF ERROR

All reasonable attempts have been taken to ensure the accuracy of the results of the longitudinal dataset. Nevertheless potential sources of error including sampling, linking and census quality error should be kept in mind when interpreting the results.

Sampling Error

Sampling error occurs because only a small proportion of the total population is used to produce estimates that represent the whole population. Sampling error refers to the fact that for a given sample size, each sample will produce different results, which will usually not be equal to the population value.

There are two common ways of reducing sampling error - increasing sample size and utilising an appropriate selection method (for example, multi-stage sampling would be appropriate for household surveys). Given the large sample size for the ACLD (1 in 20 persons), and simple random selection, sampling error is minimal.

Linking Accuracy

False links can occur during the linkage process as even when a record pair matches on all or most linking fields, it may not actually belong to the same individual. While the methodology is designed to ensure that the vast majority of links are true, some are nevertheless false. The nature of the process used for the ACLD linkage means that while the links obtained are to a high degree of accuracy, some false links may be present within the ACLD dataset. There is an estimated 5–10% false link rate in the ACLD.

For further detail on the accuracy of the linkage, see Australian Census Longitudinal Dataset, Methodology and Quality Assessment (cat. no. 2080.5).

Managing Census Quality

The ABS aims to produce high quality data from the Census. To achieve this, extensive effort is put into Census form design, collection procedures, and processing procedures.

There are four principle sources of error in Census data: respondent error, processing error, partial response and undercount. Quality management of the Census program aims to reduce error as much as possible, and to provide a measure of the remaining error to data users, to allow them to use the data in an informed way.

Respondent error

For most households in Australia, the Census is self-enumerated. This means that householders are required to complete the Census form themselves, rather than having the help of a Census collector. The Census form may be completed by one household member on behalf of others. Error can be introduced if the respondent does not understand the question, or does not know the correct information about other household members. Self-enumeration carries the risk that wrong answers could be given, either intentionally or unintentionally.

Processing Error

Much of the data on the Census form is recorded using automatic processes, such as scanning, Intelligent Character Recognition and other automatic processes. Quality assurance procedures are used during Census processing to ensure processing errors are kept at an acceptable level. Sample checking is undertaken during coding operations, and corrections are made where necessary.

Partial Response

When completing their Census form, some people do not answer all the questions which

apply to them. While questions of a sensitive nature are generally excluded from the Census, all topics have a level of non-response. This can be measured and is generally low. In those instances where a householder fails to answer a question, a 'not stated' code is allocated during processing, with the exception of non-response to age, sex, marital status and place of usual residence. These variables are needed for population estimates, so they are imputed using other information on the Census form, as well as information from the previous Census.

Undercount

The goal of the Census is to obtain a complete measure of the number and characteristics of people in Australia on Census Night and their dwellings, but it is inevitable that a small number will be missed and some will be counted more than once. In Australia more people are missed from the Census than are counted more than once, thus the effect when both factors are taken into account is a net undercount.

For more detail see Managing Census Quality.

DATA CONSISTENCY

A small percentage of linked records have inconsistent data, such as a different country of birth at the two time points or an age inconsistency of more than one year (when the expected five year difference is accounted for). Inconsistencies may be due to:

- false links - the record pair does not belong to the same individual.
- reporting error - information for the same individual was reported differently in 2006 and 2011.
- processing error - the value of a data item was inaccurately assigned during processing.

In most analysis, the effect of inconsistent information may only have a small impact. Characteristics from either the 2006 or 2011 data can be used in tables and some exploration of consistency over time will assist in drawing appropriate conclusions.

No data editing was applied to the file beyond that which had already taken place during the relevant Census processing period. A set of consistency flags has been included on the ACLD file so that inconsistent data may be observed, quantified or excluded from calculations. Consistency flags, located in the Longitudinal group of data items, have been created for Census variables that would not be expected to change over time or have unlikely transitions over time. These are as follows:

- Age
- Birthplace of Person
- Birthplace of Male Parent
- Birthplace of Female Parent
- Sex
- Year of Arrival
- Number of Children Ever Born
- Registered Marital Status
- Highest Year of School Completed
- Level of Highest Non-School Qualification
- Country of Birth of Spouse or Partner
- Age of Spouse or Partner

- Indigenous Status

There are numerous ways to define consistency. The consistency flags have fine level categories to allow users flexibility in using their own definition of consistent or inconsistent. For example, where one Census has 'not stated' for the year of arrival data item, a user can decide whether the record should be considered consistent or not. The same applies to where the response for one Census is 'not applicable'. The labels attached to each category suggesting consistency or inconsistency will assist the user in determining which records are consistent or inconsistent for their needs.

See also Longitudinal Data Items in the Data Items chapter.

INCONSISTENT REPORTING ON THE LINKED ACLD FILE, By selected characteristics

Characteristic	Proportion of linked records with inconsistent data between 2006 and 2011
	%
Age (within 1 year)	2.4
Sex	0.1
Birthplace of Person	2.1
Birthplace of Female Parent	4.0
Birthplace of Male Parent	4.4
Year of Arrival	16.5
Indigenous Status (either newly identified or previously identified as Aboriginal and/or Torres Strait Islander)	0.5
Registered Marital Status	0.7
Highest Year of School Completed	6.3
Level of Highest Non-School Qualification	14.9
Country of Birth of Spouse or Partner	2.7
Age of Spouse or Partner	7.9

The ACLD-SSRI file includes additional consistency flags with respect to the SSRI data. Due to small differences in data extraction dates for specific benefits, there are some instances of individuals having multiple records with inconsistent data in the source SSRI file. Where this occurred, one value for the inconsistent data item was selected at random from the SSRI file. These flags identify the records and data item where inconsistency occurred. For example, the AGEDUP field indicates that there were inconsistent age records for an individual SSRI recipient. In general, the effect of this inconsistency is minimal and should only significantly affect analysis results where small populations occur.

Where there are cases of inconsistency between ACLD and SSRI data, there are no consistency flags.

Where demographic information is available from both the ACLD and SSRI data it is recommended to use the Census Demographics as they will align best with most of the information on the dataset, including the weights. However there are a few cases where using the SSRI data may be more appropriate. For example, if looking at Age Pension results but using the Census age a user may observe some records that appear to be on the Age Pension despite not being Age Pension age at the time of the Census. In situations like these, it may be more appropriate to use the SSRI age for individuals where it exists.

Limitations

LIMITATIONS

While every effort has been made to assure the quality of the statistics presented in this publication, they should be considered experimental and treated with caution.

While the linkage rate of SSRI records to the ACLD is believed to be relatively high, there remain a number of records where no link was able to be made. Certain sub-groups (such as highly mobile populations) are known to have a lower propensity to link, meaning they may be under-represented in the linked dataset.

The use of weights aims to adjust for the fact that the linked records may not be representative of the whole population and correct for any under-representation of particular population sub-groups. This aims to ensure that valid inferences can be made from the linked dataset.

The weights on the ACLD-SSRI dataset reflect the benchmarking of the ACLD records only to estimates of the 2006-2011 longitudinal Australian population, that is, an estimate of the number of people resident in Australia at the times of both the 2006 and 2011 Censuses. While this approach largely accounts for the missed links between the ACLD 2006 and 2011 samples, it does not account for missed links between the ACLD and SSRI. As the overlap of ACLD and SSRI datasets is unknown, it is difficult to calculate accurate weights to account for these missed links.

As a result, caution should be used in drawing inferences from analysis of the ACLD-SSRI dataset. As the weights do not account for missed links between ACLD and SSRI records, estimates of the number of people with particular characteristics using SSRI data items may be underestimates of the actual 'real world' number of people with that characteristic. The extent of the underestimation is likely to differ for different population groups.

The dataset should therefore not be used to generate estimates of the number, or proportion, of people on a particular payment.

For example, using unlinked SSRI data we know that approximately 2.2 million Australians received the Age Pension around the time of the 2011 Census. However, the weighted estimate using the ACLD-SSRI dataset shows 2.1 million Age Pension recipients at the same time point. The difference could be due to either (1) missed links between ACLD and SSRI records, (2) differences between the longitudinal Australian population and the point in time SSRI population, or (3) a combination of (1) and (2). Similarly, using unlinked SSRI data we know that approximately 510,000 Australians received Newstart Allowance around the time of the 2011 Census, compared with a weighted longitudinal estimate of only 360,000 using the ACLD-SSRI dataset.

Analysis showing the proportion of people on a payment is likely to be an underestimate, and the extent of this underestimate may also vary between different groups.

However, the dataset does provide valuable information on the characteristics of those people receiving a benefit, both at the time of receipt of the benefit and five years previously. While it should not be used to generate proportion of people on a particular payment, the dataset should be used to examine the distribution of people on a particular payment with various characteristics. For example, the proportion of people on Newstart Allowance that volunteered for an organisation or group, or the proportion of people on the Age Pension living in a nursing home. It should be noted though that these statistics may also be subject

to some bias due to missed links.

While the majority of people resident in Australia in 2011 were also resident in 2006, as the SSRI data was linked to a dataset reflecting the 2006–2011 longitudinal Australian population, results may not be representative of the point in time population in 2011. In particular, recent migrants to Australia and new births since the 2006 Census are not in scope of the ACLD. Cross-sectional estimates should be interpreted with caution.

Due to the missed links, one cannot know with certainty for any given record whether that person was not in receipt of a benefit, or if they were in receipt of a benefit but just were not able to link. Consequently, one should not conclude that these records did not receive a payment. Any comparisons should be to the total population rather than the residual category (i.e. persons on a payment compared with the total population, rather than those on a payment compared with those not on a payment).

These same limitations apply to the use of migrant-related data linked to the ACLD. Therefore, analysis of those records that have both migrant-related information and SSRI information should be performed with particular caution.

File structure

FILE STRUCTURE

The ACLD-SSRI product in both TableBuilder and DataLab is a single-level file that counts persons and has data items for 2006 and 2011.

In TableBuilder, data items have first been separated by year of collection. Each data item for the person in 2006 has a corresponding item in 2011. In addition, data items that provide longitudinal information derived from both years, such as the consistency in reporting of certain data items between 2006 and 2011, is in a folder labelled 'Longitudinal' (see also Using the ACLD-SSRI in TableBuilder). In addition to Census variables, there are three data items relating to migrants that have been made available via the Settlement Database (SDB) from the Department of Social Services and 44 data items relating to social security payments that have been made available via the Social Security and Related Information (SSRI) dataset from the Department of Social Services. These data items are only available for 2011.

For each year, data items are further divided into the following groups:

Geography - geographic classifications on both a usual residence and place of enumeration basis.

Person - person-level characteristics, covering demographic, culture and language, education, employment and unpaid work topics.

Dwelling Related - derived from the characteristics of the dwelling in which the person was enumerated - e.g. type of dwelling.

Household Related - derived from the characteristics of the household in which the person lives, provided the person was counted at home on Census Night - e.g. total weekly household income.

Family Related - derived from the characteristics of the family in which the person lives, provided the person was counted at home on Census Night - e.g. age of youngest child in family.

Male Parent - derived from information about the person's male parent, provided the parent

was counted in the same family on Census Night - e.g. labour force status of male parent.

Female Parent - derived from information about the person's female parent, provided the parent was counted in the same family on Census Night - e.g. labour force status of female parent.

Spouse/Partner - derived from information about the person's spouse/partner, provided the spouse/partner was counted in the same family on Census Night - e.g. highest educational attainment of spouse/partner.

Where available, migrants data were provided for the spouse/partner, male parent and/or female parent, and SSRI data for the spouse/partner.

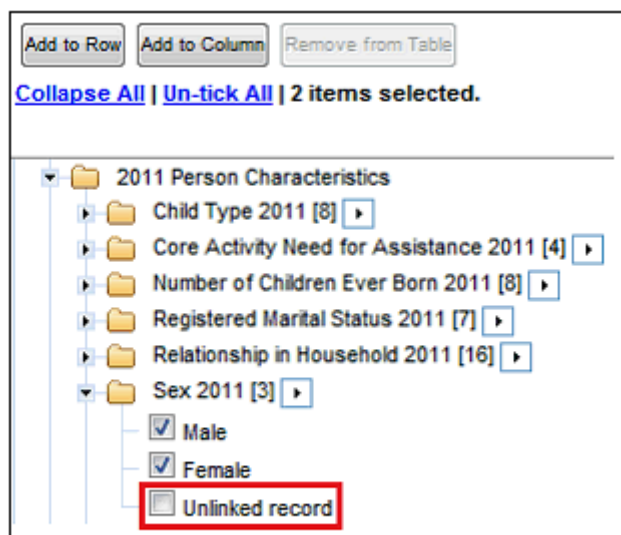
UNLINKED RECORDS

The ACLD-SSRI file contains all records from the ACLD 2006 Census sample, both those that were linked to a 2011 Census record and those that were not. Records that were linked have information for 2006 and 2011. Records that were not linked have information for 2006 only. This may assist in exploring the differences in linkage rate between the 2006 Census sample and 2011 Census records for different sub-groups.

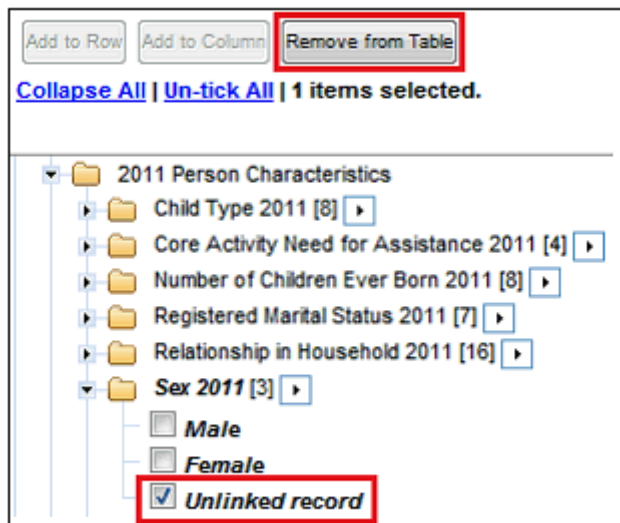
SDB and SSRI records which were not linked to ACLD records are not included in the linked dataset.

Excluding Unlinked Records in TableBuilder

When using the weighted summation option in TableBuilder, no results will be returned for unlinked records in 2011 as weights were not applied to these records. Results will be returned if analysis is performed on unweighted data only. If desired, this category can be excluded from tables by ensuring that the 'unlinked record' category is deselected before each data item is added to the table. Such a table would produce a sample count corresponding to the equivalent table run with weights. Refer to the TableBuilder, User Guide (cat. no. 1406.0.55.005) for more information on how to select data items for tables.



If the 'unlinked record' category is present on a data item that has already been added to a table, it can be removed by selecting this category within the relevant data item and then pressing the 'Remove from Table' button.



Note that removing any category, such as the 'unlinked record' category, from a table where data has already been generated will clear all data, meaning the table will need to be rerun.

Also note that the 'unlinked record' flag only refers to not achieving a link from a 2006 Census record to a 2011 Census record, without reference to the SDB or SSRI linking processes.

Using the ACLD-SSRI in TableBuilder

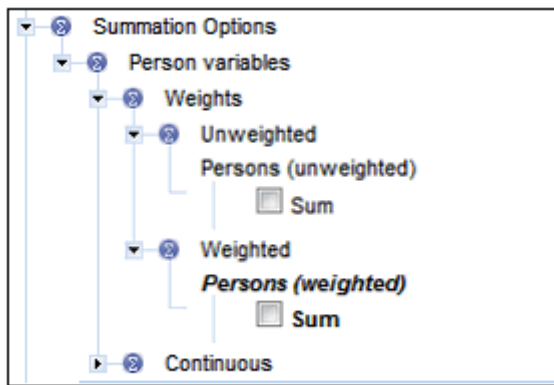
TABLEBUILDER USER GUIDE

The TableBuilder, User Guide (cat. no. 1406.0.55.005) is a comprehensive reference guide for the web interface of TableBuilder. It includes information on Getting started, Customised data and Interpreting results from ABS data.

COUNTING UNITS AND WEIGHTS

Weighting is the process of adjusting results from a sample to infer results for the total population. To do this, a weight is allocated to each person. The weight is the value that indicates how many population units are represented by the sample unit.

Both the sample and weighted count options have been made available for the ACLD-SSRI. It is therefore critical that weighted or unweighted counts are selected as appropriate when specifying tables. The following image shows the available Summation Options.



The default option used for the ACLD-SSRI is weighted count. Weights should be used when making inferences about the longitudinal Australian population and will be the basis for most analyses. Uses for unweighted counts are generally limited to research into unlinked records and more sophisticated analysis for those seeking to understand the weighting methodology better or wishing to apply their own weighting methods.

RELATIVE STANDARD ERROR

While weighted counts are available in the ACLD-SSRI TableBuilder, the Relative Standard Error will not be calculated for these counts due to the confounding effects of linking error present in the sample, which were not able to be quantified.

CONFIDENTIALITY FEATURES IN TABLEBUILDER

In accordance with the Census and Statistics Act 1905, all the data in TableBuilder are subjected to a confidentiality process before release. This confidentiality process is undertaken to avoid releasing information that may allow the identification of particular individuals, families, households, dwellings or businesses.

Processes used in the ACLD-SSRI in TableBuilder to confidentialise records include the following:

- perturbation of data
- table suppression

Perturbation of data

To minimise the risk of identifying individuals in aggregate statistics, a technique is used to randomly adjust cell values. This technique is called perturbation. Perturbation involves small random adjustments of the statistics and is considered the most satisfactory technique for avoiding the release of identifiable statistics while maximising the range of information that can be released. These adjustments have a negligible impact on the underlying pattern of the statistics.

The introduction of these random adjustments results in tables not adding up. Randomly adjusted individual cells will be consistent across tables, but the totals in any table will not be the sum of the individual cell values. The size of the difference between summed cells and the relevant total will generally be very small, as demonstrated below.

Persons (unweighted)	
Persons (unweighted) (x1)	
Sex 2011 ↑↓⊕⊗	↑↓
Male	390,467.0
Female	410,293.0
Total	800,759.0

(Sum of cells = 390,467 + 410,293 = 800,760. Difference of 1 relative to displayed total.)

Table suppression

Some tables generated within TableBuilder may contain a substantial proportion of very low counts within cells (excluding cells that have counts of zero). When this occurs, all values within the table are suppressed in order to preserve confidentiality. The following error message displayed at the bottom of the table indicates when table suppression has occurred.

'ERROR: The table has been suppressed as it is too sparse'.

The ACLD-SSRI in ABS DataLab

THE ACLD-SSRI IN ABS DATALAB

The DataLab is an interactive data analysis solution available for high end users to run advanced multivariate statistical analyses, for example multiple regressions and structural equation modelling, using SPSS20, Stata13 and SAS9. Controls in the DataLab have been put in place to protect the identification of individuals and organisations. These controls include environmental protections, data de-identification and confidentialisation, access safe guards and output clearance. All output from DataLab sessions is cleared by an ABS officer before it is released.

For more information about the DataLab please see About the DataLab. To apply for access to the ACLD or ACLD-SSRI in the DataLab, please contact Microdata Access Strategies via microdata.access@abs.gov.au.

For more information about accessing the ACLD test file for DataLab, please see ABS DataLab ACLD-SSRI Test File.

ABS DataLab ACLD-SSRI test file

ABS DATALAB ACLD-SSRI TEST FILE

The test file does not contain real data, and cannot be used for analysis.

A test file has been created for the ACLD-SSRI microdata product. The purpose of the test file is to allow researchers/analysts to become familiar with the data structure and prepare code/programs prior to applying for, or commencing, a DataLab session. This aims to

maximise the value of sessions by saving users' time and resources once they enter the DataLab environment.

The test file mimics the structure of the ACLD-SSRI microdata - it has the same data items and allowed values, however, all data in the test file is false, created through a randomisation process. Proportions of values within data items in the test file will be similar to those in the real data; however, relationships between data items are not (intentionally) maintained. It is extremely unlikely that a record in the test file would match with a genuine record in the real data.

The test file is available as a free download from the Downloads tab. It can also be made available in other file formats on request, if required. For further information users should email microdata.access@abs.gov.au.

Data items

DATA ITEMS LIST

A complete list of data items included on the ACLD-SSRI file is provided in an Excel spreadsheet that can be accessed from the Downloads tab.

All data items are created at the person level. This includes data items relating to the family and household of the person selected in the sample. For ease of use, these data items have been divided into Person, Dwelling, Household, Family, Spouse Related, and Male and Female parent related groupings.

Users intending to subscribe to the TableBuilder product or use the microdata product in the ABS Data Laboratory should ensure the data they require, and the level of detail required, are available and applicable for the intended use.

The ACLD has been updated since its initial release to add several new data items along with improved derivations for select data items. For the 12 February 2016 update certain data items were more finely disaggregated to expand analytical possibilities. For example, religious affiliation was expanded to the three digit level and year of arrival in Australia in single years. In addition to new Census variables, such as Socio-Economic Indexes of Areas (SEIFA) and dwelling location, three new data items were added from the Department of Social Services' (DSS) Settlement Database that have been previously made available on the Australian Census Migrants Integrated Dataset (ACMID) TableBuilder file. To further enhance the file, a new weight replaced the previous weight to better account for net overseas migration between 2006 and 2011.

The Social Security and Related Information (SSRI) data provided by DSS was added to the file in August 2017, creating the ACLD-SSRI, and includes demographic information (such as Age, Sex, Country of Birth, Partner Status, Marital Status, and Indigenous Indicator), location information and Benefit flags for both the individual and partners, where available. Where data was available, details of an individual's current partner as of the SSRI extract period was included. This includes the details of any benefits the partner was receiving as well as location and demographic details of the partner. Note that demographic, location and partner information from the SSRI dataset may be different to that reported in the linked Census record.

Benefits have been split into current and suspended flags for each of the benefits as an

individual may be claiming multiple benefits. An individual is considered suspended where they are temporarily ineligible for benefits due to not meeting one of the income, assets, or activity tests or suspended pending provision of additional documents. Numeric derived fields for the number of benefits an individual is receiving have also been included. Supplementary benefits, such as Rent Assistance, have not been included in this data. Monetary values related to benefit payments have not been included in the dataset.

Due to small numbers of people on particular payments, some benefits have been grouped.

VISITORS ON CENSUS NIGHT

Overseas visitors were excluded from the 2006 ACLD sample. The ACLD and ACLD-SSRI, however, do include visitors from within Australia in 2006 and/or 2011. These are people who were enumerated away from their usual residence on Census Night. Family information cannot be derived for these persons and as such all family, spouse, and male and female parent related data items are not applicable for visitors.

All dwelling related data items, however, have been made applicable to visitors. This information relates to their dwelling of enumeration on Census Night, not usual residence.

Most household data items are not applicable to visitors, however for four data items, visitors have been included in order to align to standard Census derivations of that data item. These comprise:

- Total Household Income as stated (weekly) of household in which person was enumerated.
- Total Household Income (weekly) of household in which person was enumerated.
- Household Income Derivation Indicator of household in which person was enumerated.
- Household Composition of household in which person was enumerated.

Any applicable household information for a visitor relates to their place of enumeration, not usual residence.

Where a data item is also applicable to visitors, the usual address indicator data item for the relevant Census year can be used to restrict the table to usual residents only.

The cell comments available in the data item list provide precise information on who is, and is not, applicable for each data item.

PERSONS TEMPORARILY ABSENT ON CENSUS NIGHT

The Census household form provides the opportunity to list up to three persons who were temporarily absent from the dwelling on Census Night. A limited amount of information is collected for these persons and it is used to better derive the family and household characteristics of the dwelling. In deriving family and household related data items for the ACLD and ACLD-SSRI, information on persons temporarily absent was included where relevant and available. Details are provided in cell comments in the data items list.

NOT APPLICABLE CATEGORIES

Most data items in the ACLD-SSRI include a 'not applicable' category. The definition of the 'not applicable' category, where relevant, can be found in the Census Dictionary Glossary and A guide to Australian government payments.

NOT STATED CATEGORIES

'Not stated' categories occur when no response has been provided for a data item. All Census data items contain 'not stated' categories except for age, sex, marital status and usual address, as this information is imputed for these items.

Other Census products commonly use the symbol '&' to denote a code value of 'not stated'. In the ACLD-SSRI, the symbol 'X' or '97' has been used. The codes are listed in the data items list.

Conditions of use

USER RESPONSIBILITIES

The Census and Statistics Act 1905 includes a legislative guarantee to respondents that their confidentiality will be protected. This is fundamental to the trust that the Australian public has in the ABS and that trust is in turn fundamental to the excellent quality of ABS information. Without that trust, survey respondents may be less forthcoming or truthful in answering our questionnaires. For more information, see 'Avoiding inadvertent disclosure' and 'Microdata' in How the ABS keeps your information confidential.

TABLEBUILDER

In accordance with the Census and Statistics Act 1905, data in TableBuilder are subjected to a confidentiality process before release. The release of microdata must satisfy the ABS legislative obligation to release information in a manner that is not likely to enable the identification of a particular person or organisation. This confidentiality process is applied to avoid releasing information that may lead to the identification of individuals, families, households, dwellings or businesses.

Prior to being granted access to TableBuilder users must agree to the following ABS Terms and Conditions of TableBuilder Access:

- understand that the ABS has taken great care to ensure that the information on the survey output record file is correct and as accurate as possible and understand that ABS does not guarantee, or accept any legal liability whatsoever arising from, or connected to, the use of any material contained within, or derived from TableBuilder.
- understand that all data extracted from the dataset through TableBuilder will be confidentialised prior to being supplied and that as a result, no reliance should be placed on small cells as they are impacted by random adjustment, and respondent and processing errors.
- users inform the ABS, through their Contact Officer, upon leaving their organisation so that their access is disabled.
- not to provide their TableBuilder user ID and password access to any other

person or organisation.

DATALAB

The DataLab service is provided on a fully cost recovered basis, with a one-off set up cost, and a per session cost. The costs includes preparation of the desktop PC, maintenance of the client's access to a microdata product, as well as vetting of all output produced during the DataLab session. Sessions are 3.5 hours or part thereof. Loading of different analytical packages may incur additional costs.

Access is provided to visiting users in a secure location within ABS premises. DataLab allows interactive (real time) access to the microdata files. There are no limitations on what unit record or summary information a user can view within DataLab, and users are allowed to produce graphical outputs. Users are supervised at all times and must not bring mobile phones, cameras, USB keys, laptops, palm pilots or similar transmission or storage devices into the secure location. All outputs produced by users in DataLab are manually cleared for release after the session.

For more information on DataLab, please refer the About the DataLab page on the ABS website.

CONDITIONS OF SALE

All ABS products and services are provided subject to the ABS Disclaimer, ABS Copyright, ABS Privacy and ABS Conditions of Sale.

Any queries relating to these Conditions of Sale should be emailed to intermediary.management@abs.gov.au.

PRICE

Microdata access is priced according to the ABS Pricing Policy and Commonwealth Cost Recovery Guidelines. For details refer to ABS Pricing Policy on the ABS website. For microdata prices refer to the Microdata prices web page.

HOW TO APPLY FOR ACCESS

Clients wishing to access the microdata should read the How to Apply for Microdata web page.

Clients should familiarise themselves with the Responsible Use of ABS Microdata, User Guide (cat. no. 1406.0.55.003) and/or the TableBuilder, User Guide (cat. no. 1406.0.55.005) depending on the product or products being accessed.

Other related microdata information is available via the Microdata web pages.

To apply for access to the ACLD or ACLD-SSRI in TableBuilder, register and apply in the Registration Centre.

To apply for access to the ACLD or ACLD-SSRI in DataLab, please contact Microdata

Access Strategies via microdata.access@abs.gov.au.

Further information on access steps can be found on the [How to Apply for Microdata](#) page on the ABS web site.

AUSTRALIAN UNIVERSITIES

The ABS/Universities Australia Agreement provides participating universities with access to a range of ABS products and services. This includes access to microdata.

For further information, university clients should refer to the [ABS/Universities Australia Agreement](#) web page.

FURTHER INFORMATION

The Microdata Entry page on the ABS website contains links to microdata related information to assist users to understand and access microdata.

For further information users should email microdata.access@abs.gov.au or telephone (02) 6252 7714.

Future considerations

FUTURE CONSIDERATIONS

The ACLD-SSRI dataset is a unique data source in that it brings together information about the characteristics, circumstances, and transitions of people who have interacted with the social security system, and therefore has the potential to increase knowledge about a wide range of socio-economic issues facing Australians and their families.

While some survey collections (such as the Survey of Income and Housing and the General Social Survey) include items about government payments received combined with a range of socio-economic and demographic characteristics, the substantially larger sample size of the ACLD-SSRI dataset allows these issues to be explored in much more detail. In addition, the use of administrative data gives us information on the actual payments received, while traditional surveys rely on the respondent accurately reporting. Additionally, the ACLD-SSRI allows longitudinal characteristics of benefit recipients to be explored.

However, the ACLD-SSRI dataset has some significant limitations (see [Limitations](#)) and is considered experimental in nature. Evaluation of the project to date has identified areas where different methodologies could be considered in the future to increase the robustness of the data, and where further enhancements could be made to improve the utility of the dataset.

Linkage strategy

Linkage of the ACLD (a random 5% sample of 2006 Census records linked to records from the 2011 Census) to a subset of the SSRI dataset (people who received social security benefits or who had suspended benefits in September 2011) was undertaken using selected variables common to both datasets to ensure each record in the ACLD had the highest

possible chance of being accurately linked to a record in the SSRI dataset (refer to Methodology for more information). However, the overlap of these datasets is unknown, so it is difficult to calculate an exact linkage rate.

This method was chosen as a non-identifying grouped numeric code is available only for the records in the 5% ACLD sample, and not for the full Census dataset. It was thought that the inclusion of a non-identifying grouped numeric code as a linkage variable would produce a higher quality linkage than linking to the full 2011 Census dataset where such a code is not available. However, linkage to the full Census has a number of advantages:

- Analysis of the linked ACLD-SSRI dataset has shown that most items of analytical interest are cross-sectional in nature - examining the characteristics of SSRI recipients using data items from the Census that are not available on SSRI data in isolation. Linkage to the Census itself rather than a longitudinal sample would be more optimal for this purpose (the relevant ACLD sample could then be extracted from the linked dataset for any required longitudinal analysis).
- Linkage to the 2011 Census allows a direct calculation of a linkage rate as a measure of linkage quality.
- While the ACLD-SSRI dataset is large in comparison to sample surveys, linkage to the full Census dataset allows the characteristics of very small groups or small areas to be examined.

Weighting

The ACLD-SSRI record weights are not benchmarked to the SSRI population. While the weights are of good quality for use in analysing the longitudinal Australian population, there is no mathematical measure of how accurately they reflect the demographic make-up of either the longitudinal SSRI population or the "point in time" population at the time of the Census. In future, alternative linking strategies may allow for good quality benchmarks incorporating SSRI population information to be produced, enabling the calculation of more accurate weights, thus increasing the utility of the file and allowing inferences to be made with more confidence.

Weighting the linked dataset to be representative of the input SSRI dataset would minimise the impact of bias on analysis of the characteristics of SSRI recipients.

Additional data items

This pilot project involved linking SSRI data from the 2011 time-point only. Analysis has shown that it may have been more analytically useful to include SSRI data from earlier time points in order to use 2011 Census data to explore the outcomes for those on particular payments. Longitudinal benefits information would provide a valuable resource for understanding the dynamics and pathways of people's involvement in the social security system.

For the purposes of this pilot project, only a relatively small selection of SSRI data items was linked to the ACLD dataset. While this has resulted in a unique and valuable dataset bringing together information not available on any alternative dataset, there may be potential in the future to further enhance the dataset by including a greater range of SSRI data, or information from additional time-points.

About this Release

Information from the Department of Social Services' (DSS) Social Security and Related Information (SSRI) dataset was added to the 2006-2011 Australian Census Longitudinal Dataset (ACLD) to create a new dataset (ACLD-SSRI). The SSRI data provided to ABS was securely linked to the ACLD in accordance with the High Level Principles for Statistical Data Integration and is protected by the secrecy provisions of the Census and Statistics Act 1905. The linkage of the ACLD with the SSRI data provides valuable information on the characteristics of people receiving a benefit, both at the time of receipt of the benefit and five years previously. Includes microdata and aggregate data.

Explanatory Notes

Glossary

Terms and definitions applicable to the ACLD-SSRI can be located via the Census Dictionary Glossary and A guide to Australian government payments.

Abbreviations

ABS	Australian Bureau of Statistics
ABSDL	Australian Bureau of Statistics Data Laboratory
ACLD	Australian Census Longitudinal Dataset
ACMID	Australian Census Migrants Integrated Dataset
ASGS	Australian Statistical Geography Standard
DHS	Department of Human Services
DSS	Department of Social Services
ERP	Estimated Resident Population
SA2	Statistical Area 2
SDB	Settlement Database
SSRI	Social Security and Related Information

Quality Declaration

INSTITUTIONAL ENVIRONMENT
RELEVANCE
TIMELINESS
ACCURACY
COHERENCE
INTERPRETABILITY
ACCESSIBILITY

INSTITUTIONAL ENVIRONMENT

The Australian Census Longitudinal Dataset with Social Security and Related Information (ACLD-SSRI) is released in TableBuilder and as a microdata product in DataLab. Microdata files are released in accordance with the conditions specified in the Statistics Determination section of the Census and Statistics Act 1905. This ensures that confidentiality is maintained whilst enabling micro level data to be released. More information on the confidentiality practices associated with TableBuilder, User Guide (cat. no. 1406.0.55.005) can be found on the Confidentiality page. To protect confidentiality of data within DataLab, users are supervised at all times and must not bring mobile phones, cameras, USB keys, laptops, palm pilots or similar transmission or storage devices into the secure location. All outputs produced by users in DataLab are manually cleared for release after the session.

For information on the institutional environment of the Australian Bureau of Statistics (ABS), including the legislative obligations of the ABS, financing and governance arrangements, and mechanisms for scrutiny of ABS operations, see ABS Institutional Environment.

In April 2012, the ABS became an accredited Integrating Authority under the Commonwealth data integration interim arrangements. A copy of the accreditation claims made by the ABS, which have been verified by an independent auditor, is available through the National Statistical Service (NSS) website. The ABS only undertakes data integration for statistical and research purposes and where there is a strong public benefit in doing so.

RELEVANCE

Data for the Census of Population and Housing used in this product were collected on 8 August 2006 and 9 August 2011. The scope of the Census is all persons in Australia on Census night excluding certain diplomatic personnel of overseas governments. In addition persons who were overseas visitors on Census Night were excluded from the scope of the 2006 ACLD sample.

The Census collects information on demographics, income, labour force, unpaid work, dwelling characteristics and family and household relationships.

For more information, see How Australia Takes a Census, 2011 (cat. no. 2903.0).

Settlement Database (SDB) data contains information on persons who had a permanent visa with a date of arrival between 1 January 2000 and 9 August 2011 (inclusive). The SDB date of arrival on which the scope is based reflects an individual's latest arrival pertaining to their latest permanent visa. For an offshore applicant, the SDB arrival date is when the applicant arrives in Australia on that permanent visa. However, for a person who applies onshore for a permanent visa, the date of arrival listed on the SDB is the date of their last entry into Australia.

SSRI data contains information on persons receiving selected Australian Government benefits as of the end of September 2011. It includes those persons who were on suspended benefits at this time, as well as their partner information where available.

For further information, particularly regarding analysis limitations of the ACLD-SSRI records with both migrants and SSRI information, refer to Scope and Coverage in the Methodology section.

TIMELINESS

The Census of Population and Housing is conducted every five years. For further

information see the publication *How Australia Takes a Census, 2011* (cat. no. 2903.0).

The first wave of Census data for the ACLD was from 2006 and the second from 2011. SDB data covers visa holders whose last arrival was made between 1 January 2000 and 9 August 2011 (inclusive) and SSRI data was from September 2011.

ACCURACY

The ACLD was created using data linkage techniques without name and address but with other characteristics from the Census. It was based on a 5% random sample from the 2006 Census (979,661 records) of which 82% (800,759) were linked to a 2011 Census record. The false link rate is estimated at around 5–10%. In the linkage of 2011 Census records to SDB records, a 76% linkage rate was achieved, see *Microdata: Australian Census and Migrants Integrated Dataset, 2011* (cat.no. 3417.0.55.001). In the linkage of ACLD 2011 Census records to SSRI records, as the overlap of these two datasets is unknown, it is difficult to calculate an exact linkage rate for this exercise, see *Linking Methodology* in Methodology section of this publication.

Sampling error occurs because only a small proportion of the total population is used to produce estimates that represent the whole population. Sampling error refers to the fact that for a given sample size, each sample will produce different results, which will usually not be equal to the population value. There are two common ways of reducing sampling error - increasing sample size and utilising an appropriate selection method (for example, multi-stage sampling would be appropriate for household surveys). Given the large sample size for the ACLD (1 in 20 persons), and simple random selection, sampling error is minimal.

The ACLD sample was weighted to an estimate of the population that was resident in Australia at the time of both the 2006 and 2011 Censuses. The weights adjust for missed links and Census undercount.

Information on methodology, linkage quality and weighting can be found in *Information Paper: Australian Census Longitudinal Dataset - Methodology and Quality Assessment, 2006 and 2011* (cat. no. 2080.5). Information on the weighting of the ACLD has since changed with the update of the ACLD file that was released on 11 February, 2016. For information on the current weighting please see the Methodology section of this publication.

The ACLD weights are not benchmarked to the SDB or SSRI populations, so weighted estimates for SDB and SSRI populations should be interpreted with caution.

Steps are taken to confidentialise the data made available on TableBuilder in such a way as to maximise the usefulness of the content while maintaining the confidentiality of respondents selected in the ACLD sample. As a result it may not be possible to exactly reconcile all the statistics produced from the microdata with other published statistics. Further information about the steps taken to confidentialise the microdata can be found in TableBuilder confidentiality.

Estimates of migrants in the ACLD-SSRI using the SDB information may differ from the estimates produced in the ACMID for several reasons:

- the ACLD-SSRI is a 5% random sample from the 2006 Census population, not the 2011 Census population;
- more recent migrants, that is, those that arrived between the 2006 and 2011 Censuses, are not included in the sample;
- linked records in the ACLD-SSRI were benchmarked to an estimate of the

population that was resident in Australia at the time of both the 2006 and 2011 Censuses;

- the SDB information was merged to the 2011 Census records of people that were able to be linked (and not all records were able to be linked).

Due to these reasons above and other quality issues in the ACLD, estimates of migrants using the SDB information should generally be treated with caution.

COHERENCE

A small percentage of linked records have inconsistent data, such as a different country of birth at the two time points or an age inconsistency of more than one year. Inconsistencies may be due to:

- false link - the record pair does not belong to the same individual
- reporting error - information for the same individual was reported differently in 2006 and 2011
- processing error - the value of a data item was inaccurately assigned or imputed during processing. ACLD-SSRI microdata contains a large number of data items and in some cases the level of detail has been collapsed from that described in the Census Dictionary. For more information on the level of detail provided, please see the associated data item list.

While the 2006 and 2011 Censuses had predominantly the same questions and were processed in a similar way, there were some differences between them.

Notable data items that are different between Census years are personal, family and household income. Income was collected in ranges and these ranges are different in different Census years. The ACLD does not include an adjustment to income data for inflation.

Some data items were derived differently between Censuses. To aid comparability, the 2006 variable was re-derived to make it consistent with the 2011 derivation.

2006 Geography was originally disseminated according to the Australian Statistical Geography Classification. There were significant changes in the geographical classification used in the Census between 2006 and 2011 with the introduction of the Australian Statistical Geographic Standard. In order to facilitate comparison, a 'best fit' of 2006 collection districts to align with the 2011 geographic boundaries was created, however in some cases this will be an approximation only.

The exception is Remoteness Areas where no adjustment was made and 2006 boundaries and definitions were applied for usual address for that time period.

For more information on the differences between the 2006 and 2011 Census see What's new for 2011.

INTERPRETABILITY

This publication should be referred to when using the microdata. It contains information on the Sample Methodology, File Structure, Using the ACLD-SSRI in TableBuilder, The ACLD-

SSRI in DataLab, Conditions of use and the Data items list.

Weights on the ACLD-SSRI attempt to account for missed links and Census undercount between the 2006 and 2011 Censuses. Weights do not attempt to account for missed links between the ACLD and either the migrant-related information or the social security information. Estimates and proportions of these groups will therefore be underestimates.

Detailed information on methodology, linkage quality and weighting can be found in Information Paper: Australian Census Longitudinal Dataset - Methodology and Quality Assessment, 2006 and 2011 (cat. no. 2080.5). Information on the weighting of the ACLD has since changed with the update of the ACLD file that was released on 11 February, 2016. For information on the current weighting please see the Methodology section of this publication.

The ABS publishes extensive information on Census Data Quality.

ACCESSIBILITY

ACLD-SSRI can be accessed through TableBuilder and DataLab.

This microdata product is available to approved users. Users wishing to access the microdata should read the How to apply for Microdata web page, before applying for access through the Registration Centre. Users should also familiarise themselves with information available via the Microdata web pages.

Any questions regarding access to microdata can be forwarded to microdata.access@abs.gov.au or phone (02) 6252 7714.